



EL 7418325 (S) 3-02
#4725
Priority Papers
INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ



I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Evans

Dated

26 January 2001

THIS PAGE BLANK (USPTO)



1/77

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

THE PATENT OFFICE
L
25 NOV 2000
NEWPORT

The Patent Office

Cardiff Road
Newport
South Wales
NP10 8QQ

1. Your reference

30010513 GB

2. Patent application number

(The Patent Office will fill in this part)

0028804.3

25 NOV 2000

3. Full name, address and postcode of the or of each applicant (underline all surnames)

Hewlett-Packard Company
3000 Hanover Street
Palo Alto
CA 94304, USA

Patents ADP number (if you know it)

496588004
Delaware, USA

If the applicant is a corporate body, give the country/state of its incorporation

4. Title of the invention Voice communication concerning a local entity

5. Name of your agent (if you have one)

"Address for service" in the United Kingdom to which all correspondence should be sent (including the postcode)

Robert F Squibbs
Hewlett-Packard Ltd, IP Section
Filton Road
Stoke Gifford
Bristol BS34 8QZ

Patents ADP number (if you know it)

4077442003 IS

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application number

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? (Answer 'Yes' if:
a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

Yes

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form

Description

19

Claim(s)

5

Abstract

1

Drawing(s)

4 *x411*

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77)

Request for preliminary examination and search (Patents Form 9/77)

1

M.

Request for substantive examination (Patents Form 10/77)

Any other documents (please specify)

11. I/We request the grant of a patent on the basis of this application.

Signature

Robert F. Squibb

Date

24 Nov 2000

12. Name and daytime telephone number of person to contact in the United Kingdom

Tel: 0117-312-

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed, if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention, and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

- If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.
- Write your answers in capital letters using black ink or you may type them.
- If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- Once you have filled in the form you must remember to sign and date it.
- For details of the fee and ways to pay please contact the Patent Office.

Voice Communication Concerning a Local Entity

Field of the Invention

- 5 The present invention relates to voice services and in particular, but not exclusively, to a method of providing for voice interaction with a local dumb device.

Background of the Invention

- In recent years there has been an explosion in the number of services available over the
 10 World Wide Web on the public internet (generally referred to as the "web"), the web being composed of a myriad of pages linked together by hyperlinks and delivered by servers on request using the HTTP protocol. Each page comprises content marked up with tags to enable the receiving application (typically a GUI browser) to render the page content in the manner intended by the page author; the markup language used for standard web pages is
 15 HTML (HyperText Markup Language).

- However, today far more people have access to a telephone than have access to a computer with an Internet connection. Sales of cellphones are outstripping PC sales so that many people have already or soon will have a phone within reach where ever they go. As a
 20 result, there is increasing interest in being able to access web-based services from phones. 'Voice Browsers' offer the promise of allowing everyone to access web-based services from any phone, making it practical to access the Web any time and any where, whether at home, on the move, or at work.

- 25 Voice browsers allow people to access the Web using speech synthesis, pre-recorded audio, and speech recognition. Figure 1 of the accompanying drawings illustrates the general role played by a voice browser. As can be seen, a voice browser is interposed between a user 2 and a voice page server 4. This server 4 holds voice service pages (text pages) that are marked-up with tags of a voice-related markup language (or languages).
 30 When a page is requested by the user 2, it is interpreted at a top level (dialog level) by a dialog manager 7 of the voice browser 3 and output intended for the user is passed in text

form to a Text-To-Speech (TTS) converter 6 which provides appropriate voice output to the user. User voice input is converted to text by speech recognition module 5 of the voice browser 3 and the dialog manager 7 determines what action is to be taken according to the received input and the directions in the original page. The voice input / output interface
 5 can be supplemented by keypads and small displays.

In general terms, therefore, a voice browser can be considered as a largely software device which interprets a voice markup language and generate a dialog with voice output, and possibly other output modalities, and / or voice input, and possibly other modalities (this
 10 definition derives from a working draft, dated September 2000, of the Voice browser Working Group of the World Wide Web Consortium).

Voice browsers may also be used together with graphical displays, keyboards, and pointing devices (e.g. a mouse) in order to produce a rich "multimodal voice browser". Voice
 15 interfaces and the keyboard, pointing device and display maybe used as alternate interfaces to the same service or could be seen as being used together to give a rich interface using all these modes combined.

Some examples of devices that allow multimodal interactions could be multimedia PC, or
 20 a communication appliance incorporating a display, keyboard, microphone and speaker/headset, an in car Voice Browser might have display and speech interfaces that could work together, or a Kiosk.

Some services may use all the modes together to provide an enhanced user experience, for
 25 example, a user could touch a street map displayed on a touch sensitive display and say "Tell me how I get here?". Some services might offer alternate interfaces allowing the user flexibility when doing different activities. For example while driving speech could be used to access services, but a passenger might used the keyboard.

Figure 2 of the accompanying drawings shows in greater detail the components of an example voice browser for handling voice pages 15 marked up with tags related to four different voice markup languages, namely:

- tags of a dialog markup language that serves to specify voice dialog behaviour;
- 5 - tags of a multimodal markup language that extends the dialog markup language to support other input modes (keyboard, mouse, etc.) and output modes (large and small screens);
- tags of a speech grammar markup language that serve to specify the grammar of user input; and
- 10 - tags of a speech synthesis markup language that serve to specify voice characteristics, types of sentences, word emphasis, etc.

When a page 15 is loaded into the voice browser, dialog manager 7 determines from the dialog tags and multimodal tags what actions are to be taken (the dialog manager being
 15 programmed to understand both the dialog and multimodal languages 19). These actions may include auxiliary functions 18 (available at any time during page processing) accessible through APIs and including such things as database lookups, user identity and validation, telephone call control etc. When speech output to the user is called for, the semantics of the output is passed, with any associated speech synthesis tags, to output
 20 channel 12 where a language generator 23 produces the final text to be rendered into speech by text-to-speech converter 6 and output to speaker 17. In the simplest case, the text to be rendered into speech is fully specified in the voice page 15 and the language generator 23 is not required for generating the final output text; however, in more complex cases, only semantic elements are passed, embedded in tags of a natural language
 25 semantics markup language (not depicted in Figure 2) that is understood by the language generator. The TTS converter 6 takes account of the speech synthesis tags when effecting text to speech conversion for which purpose it is cognisant of the speech synthesis markup language 25.

30 User voice input is received by microphone 16 and supplied to an input channel of the voice browser. Speech recogniser 5 generates text which is fed to a language understanding

module 21 to produce semantics of the input for passing to the dialog manager 7. The speech recogniser 5 and language understanding module 21 work according to specific lexicon and grammar markup language 22 and, of course, take account of any grammar tags related to the current input that appear in page 15. The semantic output to the dialog manager 7 may simply be a permitted input word or may be more complex and include embedded tags of a natural language semantics markup language. The dialog manager 7 determines what action to take next (including, for example, fetching another page) based on the received user input and the dialog tags in the current page 15.

- 10 Any multimodal tags in the voice page 15 are used to control and interpret multimodal input/output. Such input/output is enabled by an appropriate recogniser 27 in the input channel 11 and an appropriate output constructor 28 in the output channel 12.

15 Whatever its precise form, the voice browser can be located at any point between the user and the voice page server. Figures 3 to 5 illustrate three possibilities in the case where the voice browser functionality is kept all together; many other possibilities exist when the functional components of the voice browser are separated and located in different logical/physical locations.

- 20 In Figure 3, the voice browser 3 is depicted as incorporated into an end-user system 8 (such as a PC or mobile entity) associated with user 2. In this case, the voice page server 4 is connected to the voice browser 3 by any suitable data-capable bearer service extending across one or more networks 9 that serve to provide connectivity between server 4 and end-user system 8. The data-capable bearer service is only required to carry text-based pages and therefore does not require a high bandwidth.
- 25

Figure 4 shows the voice browser 3 as co-located with the voice page server 4. In this case, voice input/output is passed across a voice network 9 between the end-user system 8 and the voice browser 3 at the voice page server site. The fact that the voice service is embodied as voice pages interpreted by a voice browser is not apparent to the user or

30

network and the service could be implemented in other ways without the user or network being aware.

In Figure 5, the voice browser 3 is located in the network infrastructure between the end-
5 user system 8 and the voice page server 4, voice input and output passing between the end-
user system and voice browser over one network leg, and voice-page text data passing
between the voice page server 4 and voice browser 3 over another network leg. This
arrangement has certain advantages; in particular, by locating expensive resources (speech
recognition, TTS converter) in the network, they can be used for many different users with
10 user profiles being used to customise the voice-browser service provided to each user.

A more specific and detailed example will now be given to illustrate how voice browser
functionality can be differently located between the user and server. More particularly,
Figure 6 illustrates the provision of voice services to a mobile entity 40 which can
15 communicate over a mobile communication infrastructure with voice-based service
systems 4, 61. In this example, the mobile entity 40 communicates, using radio subsystem
42 and a phone subsystem 43, with the fixed infrastructure of a GSM PLMN (Public Land
Mobile Network) 30 to provide basic voice telephony services. In addition, the mobile
entity 40 includes a data-handling subsystem 45 interworking, via data interface 44, with
20 the radio subsystem 42 for the transmission and reception of data over a data-capable
bearer service provided by the PLMN; the data-capable bearer service enables the mobile
entity 40 to access the public Internet 60 (or other data network). The data handling
subsystem 45 supports an operating environment 46 in which applications run, the
operating environment including an appropriate communications stack.

25

Considering the Figure 6 arrangement in more detail, the fixed infrastructure 30 of the
GSM PLMN comprises one or more Base Station Subsystems (BSS) 31 and a Network
and Switching Subsystem NSS 32. Each BSS 31 comprises a Base Station Controller
(BSC) 34 controlling multiple Base Transceiver Stations (BTS) 33 each associated with a
30 respective "cell" of the radio network. When active, the radio subsystem 42 of the mobile
entity 20 communicates via a radio link with the BTS 33 of the cell in which the mobile

entity is currently located. As regards the NSS 32, this comprises one or more Mobile Switching Centers (MSC) 35 together with other elements such as Visitor Location Registers 52 and Home Location Register 52.

- 5 When the mobile entity 40 is used to make a normal telephone call, a traffic circuit for carrying digitised voice is set up through the relevant BSS 31 to the NSS 32 which is then responsible for routing the call to the target phone whether in the same PLMN or in another network such as PSTN (Public Switched Telephone Network) 56.
- 10 With respect to data transmission to/from the mobile entity 40, in the present example three different data-capable bearer services are depicted though other possibilities exist. A first data-capable bearer service is available in the form of a Circuit Switched Data (CSD) service; in this case a full traffic circuit is used for carrying data and the MSC 35 routes the circuit to an InterWorking Function IWF 54 the precise nature of which depends on what is
15 connected to the other side of the IWF. Thus, IWF could be configured to provide direct access to the public Internet 60 (that is, provide functionality similar to an IAP - Internet Access Provider IAP). Alternatively, the IWF could simply be a modem connecting to PSTN 56; in this case, Internet access can be achieved by connection across the PSTN to a standard IAP.
- 20 A second, low bandwidth, data-capable bearer service is available through use of the Short Message Service that passes data carried in signalling channel slots to an SMS unit 53 which can be arranged to provide connectivity to the public Internet 60.
- 25 A third data-capable bearer service is provided in the form of GPRS (General Packet Radio Service which enables IP (or X.25) packet data to be passed from the data handling system of the mobile entity 40, via the data interface 44, radio subsystem 41 and relevant BSS 31, to a GPRS network 37 of the PLMN 30 (and vice versa). The GPRS network 37 includes a SGSN (Serving GPRS Support Node) 38 interfacing BSC 34 with the network 37, and a
30 GGSN (Gateway GPRS Support Node) interfacing the network 37 with an external network (in this example, the public Internet 60). Full details of GPRS can be found in the

ETSI (European Telecommunications Standards Institute) GSM 03.60 specification. Using GPRS, the mobile entity 40 can exchange packet data via the BSS 31 and GPRS network 37 with entities connected to the public Internet 60.

- 5 The data connection between the PLMN 30 and the Internet 60 will generally be through a gateway 55 providing functionality such as firewall and proxy functionality.

Different data-capable bearer services to those described above may be provided, the described services being simply examples of what is possible. Indeed, whilst the above
 10 description of the connectivity of a mobile entity to resources connected to the communications infrastructure, has been given with reference to a PLMN based on GSM technology, it will be appreciated that many other cellular radio technologies exist (for example, UTMS, CDMA etc.) and can typically provide equivalent functionality to that described for the GSM PLMN 30.

15

The mobile entity 40 itself may take many different forms. For example, it could be two separate units such as a mobile phone (providing elements 42-44) and a mobile PC (providing the data-handling system 45), coupled by an appropriate link (wireline, infrared or even short range radio system such as Bluetooth). Alternatively, mobile entity 40 could
 20 be a single unit.

Figure 6 depicts both a voice page server 4 connected to the public internet 60 and a voice-based service system 61 accessible via the normal telephone links.

- 25 The voice-based service system 61 is, for example, a call center and would typically be connected to the PSTN 56 and be accessible to mobile entity 40 via PLMN 30 and PSTN 56. The system 56 could also (or alternatively) be connected directly to the PLMN though this is unlikely. The voice-based service system 61 includes interactive voice response units implemented using voice pages interpreted by a voice browser 3A. Thus a user can
 30 user mobile entity 40 to talk to the service system 61 over the voice circuits of the

telephone infrastructure; this arrangement corresponds to the situation illustrated in Figure 4 where the voice browser is co-located with the voice page server.

If, as shown, the service system 61 is also connected to the public internet 60 and is enabled to receive VoIP (Voice over IP) telephone traffic, then provided the data handling subsystem 45 of the mobile entity 40 has VoIP functionality, the user could use a data capable bearer service of the PLMN 30 of sufficient bandwidth and QoS (quality of service) to establish a VoIP call, via PLMN 30, gateway 55, and internet 60, with the service system 61.

10

With regard to access to the voice services embodied in the voice pages held by voice page server 4 connected to the public internet 60, if the data-handling subsystem of the mobile entity is equipped with a voice browser 3E, then all that the mobile entity need do to use these services is to establish a data-capable bearer connection with the voice page server 4 via the PLMN 30, gateway 55 and internet 60, this connection then being used to carry the text based request response messages between the server 61 and mobile entity 4. This corresponds to the arrangement depicted in Figure 3.

PSTN 56 can be provisioned with a voice browser 3B at internet gateway 57 access point. This enables the mobile entity to place a voice call to a number that routes the call to the voice browser and then has the latter connect to the voice page server 4 to retrieve particular voice pages. Voice browser then interprets these pages back to the mobile entity over the voice circuits of the telephone network. In a similar manner, PLMN 30 could also be provided with a voice browser at its internet gateway 55. Again, third party service providers could provide voice browser services 3D accessible over the public telephone network and connected to the internet to connect with server 4. All these arrangements are embodiments of the situation depicted in Figure 5 where the voice browser is located in the communication network infrastructure between the user end system and voice page server.

It will be appreciated that whilst the foregoing description given with respect to Figure 6 concerns the use of voice browsers in a cellular mobile network environment, voice

browsers are equally applicable to other environments with mobile or static connectivity to the user.

Voice-based services are highly attractive because of their ease of use; however, they do
5 require significant functionality to support them. For this reason, whilst it is desirable to provide voice interaction capability for many types of devices in every day use, the cost of doing so is currently prohibitive.

It is an object of the present invention to provide a method and apparatus by which entities
10 can be given a voice interface simply and at low cost.

Summary of the Invention

According to one aspect of the present invention, there is provided a method of voice
15 communication concerning a local entity wherein:

- (a) - the local entity has an associated voice service hosted on a separate server connected to a communications infrastructure;
- (b) - upon a user approaching the local entity, contact data relating to the user is passed to a receiving device that is located at or near the local entity and is connected to the
20 communications infrastructure;
- (c) - the contact data received by the receiving device is used to establish communication through the communications infrastructure between the voice service and equipment carried by the user that is in wireless connection with the communications infrastructure;
- 25 (d) - the user interacts with the voice service with the latter acting as voice proxy for the local entity.

The present invention also encompasses apparatus for implementing the foregoing method.

Brief Description of the Drawings

A method and apparatus embodying the invention, for communicating with a dumb entity, will now be described, by way of non-limiting example, with reference to the accompanying diagrammatic drawings, in which:

- 5 . **Figure 1** is a diagram illustrating the role of a voice browser;
- . **Figure 2** is a diagram showing the functional elements of a voice browser and their relationship to different types of voice markup tags;
- . **Figure 3** is a diagram showing a voice service implemented with voice browser functionality located in an end-user system;
- 10 . **Figure 4** is a diagram showing a voice service implemented with voice browser functionality co-located with a voice page server;
- . **Figure 5** is a diagram showing a voice service implemented with voice browser functionality located in a network between the end-user system and voice page server;
- 15 . **Figure 6** is a diagram of a mobile entity accessing voice services via various routes through a communications infrastructure including a PLMN, PSTN and public internet;
- . **Figure 7** is a diagram of a first embodiment of the invention involving a mobile phone for accessing a remote voice page server; and
- 20 . **Figure 8** is a diagram of a second embodiment of the invention involving a home server system.

Best Mode of Carrying Out the Invention

- 25 In the following description, voice services are described based on voice page servers serving pages with embedded voice markup tags to voice browsers. Unless otherwise indicated, the foregoing description of voice browsers, and their possible locations and access methods is to be taken as applying also to the described embodiments of the invention. Furthermore, although voice-browser based forms of voice services are
- 30 preferred, the present invention in its widest conception, is not limited to these forms of

voice service system and other suitable systems will be apparent to persons skilled in the art.

In both embodiments of the invention to be described below with references to Figures 7 and 8 respectively, a dumb entity (here a plant 71, but potentially any object, including a mobile object) is given a voice dialog capability by associating with the plant 71 a receiving device 72 for receiving user-related contact data from user-carried equipment using a short-range wireless communication system such as an infrared system or a radio-based system (for example, a Bluetooth system), or a sound-based system. The contact data enables a voice service associated with the plant to be placed in communication with the user through a communications infrastructure – the voice service thus acts as a voice dialog proxy for the plant and gives the impression to the persons using the service that they are conversing with the plant. The user-related contact data can be a telephone number or data address of the user's equipment, or it can take the form of a user identifier which is used to look up an access number or address of the user's equipment using a user database.

Considering the Figure 7 embodiment first in more detail, a user 5 is equipped with a mobile entity 40 similar to that of Figure 6 but provided with a short-range wireless transmitter 73 (such as an infrared transmitter) for sending user-related contact data to a complementary receiving device 72 located at or near the plant 71 (see arrow 75). The receiving device 72 is connected to the internet 60 by any appropriate connection (wireline or wireless). The contact data received by the receiving device 72 is used to establish contact, across the communication infrastructure formed by PLMN 30, PSTN 56 and internet 60, between the user's mobile entity 40 and a voice service provided by a voice page server 4 that is connected to the public internet (the PSTN 56 may or may not be involved in this link up). As already described with reference to Figure 6, a number of possible routes exist through the infrastructure between the mobile entity and voice page server 4 and various ways of using these routes will now be outlined that differ according to the location of the voice browser 3 used to interpret the voice pages served by the server 4, and what the receiving device 72 does with the user-related contact data it receives.

A) - The contact data is passed by the receiving device 72 to a voice browser 3 located in the communications infrastructure together with the URL of the voice service for the plant 71, this service being in the form of voice pages hosted on voice page server 4.

5 The contact data is either a telephone number associated with the phone functionality 43 of the mobile entity or a current data address for contacting the data-handling subsystem of the mobile entity. Where the contact data is a telephone number, the voice browser calls the mobile entity to set up a voice circuit with the latter; alternatively, the voice browser can use an SMS service to send the user a number to

10 call back (the advantage of this is that main call charge will be carried by the user). At the same time, the browser accesses the voice page server 4 to retrieve a first page of the voice service associated with the plant 71. This page (and any subsequent pages) are then interpreted by the voice browser with voice output being passed over the voice circuit to the phone subsystem 43 and thus to user 5, and voice input from

15 the user being returned over the same circuit to the browser. This is the arrangement depicted by the arrows 77 to 79 in Figure 7 with arrow 77 representing the initial passing of the user-related contact data and the voice service URL to the voice browser, arrow 78 depicting the exchange of request/response messages between the browser 3 and server 4, and arrow 79 representing the exchange of voice messages

20 across the voice circuit between the voice browser 3 and phone subsystem of mobile entity 40. Where the contact data is a data address, the operation is similar to that described above but now the voice browser uses a data-capable bearer service through the communication infrastructure to initiate a session with a packetised voice application (e.g. VoIP) running in the data-handling subsystem 45 of the mobile

25 entity 40 in order to exchange voice input/output with the mobile entity.

Where the voice browser sets up the voice circuit or data connection then either the user will have to have given sufficient data and authorisation for the user's account with the PLMN to be charged, or else the charge will be borne by the party

30 responsible for the voice browser or the voice service, though arrangements may

have been pre-established by these parties for charging the user at least for the call charge itself.

5 A variant on the foregoing is where the voice browser has access to user data (in particular, to an access code or number for the user's equipment) based on knowing the user's identity. In this case, the user-related contact data need only comprise the user's identity though generally a user-input authorisation code will also be required for accessing the user data. The user data can be associated with a specific voice browser with which the user is registered (in which case the browser's contact
10 information would need to form an element of the user-related contact data); alternatively, the user data could be more generally held, for example, as part of the data held on mobile subscribers by the PLMN operator in HLR 51 (Figure 6), though again user-authorisation will generally be required for the voice browser to access the information.

15

B) - The user-related contact data (in any of the forms discussed above) is passed by the receiving device 72 to the voice page server 4 which is then responsible for initiating contact with the mobile entity 40. Where the voice pages are to be interpreted by a voice browser located at the voice page server or in the communications
20 infrastructure (including any connected service system), then the voice browser passes the contact data (and, of course, its own URL) to the voice browser and matters proceed as described above in (A). Where the voice browser is located in the mobile entity 40 (an application running in the data handling subsystem 45), then the voice page server 4 can use the contact data to establish a data connection through
25 the communications infrastructure with the data-handling subsystem 45 for the transfer of voice pages to the voice browser and the receipt of text-based requests from the latter.

C) - The user-related contact data can be used by the receiving device 72 to pass the URL
30 of its voice service to the mobile entity (for example, using an SMS service or a data connection through the communications infrastructure). The mobile entity is then

responsible for connecting to the voice service, either through the intermediary of a voice browser 3 in the communications infrastructure, or directly by a data connection (in the case where the voice browser is in the mobile entity) or a voice connection (in the case where the voice browser is at the voice page server 4).

5

Where the mobile entity 40 is itself equipped with a voice browser 3 but resources (such as memory or processing power) at the mobile entity are restricted, the data connection used by the voice browser to receive voice pages can also be used to access remote resources as may be needed, including the pulling in of appropriate lexicons and grammar

10 specifications.

Generally, the user will only operate the short-range transmitter 73 when wanting to converse with an entity (plant 71). However, it would also be possible to arrange for the user's contact data to be continually transmitted; in this case, since spurious entities of no
15 interest to the user may then pick up the contact data, the voice browser 3 is preferably arranged to confirm with the user that they wish to talk to a particular voice service before communication is allowed to go ahead.

The nature of the voice service and, in particular the dialog followed, will of course,
20 depend on the nature of the dumb entity being given a voice capability. In the present case of a plant 71, the dialog may be directed at informing the user about the plant and its general needs. In fact, by associating sensors with the plant that feed information to the receiving device, the current state and needs of the plant can be passed to the voice service along with the user-related contact data. The information about the current state and needs
25 of the plant are stored by the voice service (for example, as session data either at the voice browser or voice page server) and enables the voice service output to be conditioned to the state and needs of the plant.

30 The Figure 8 embodiment concerns a restricted environment (here taken to be a home environment but potentially any other proprietary space, or office or similar) where a home

server system 80 includes a voice page server 4 and associated voice browser 3, the latter being connected to a wireless interface 82 to enable it to communicate with devices in the home over a home wireless network. In this embodiment, user-related contact data in the form of a user identity is output by a forward-facing infrared transmitter 83 mounted on a wireless headset 90 worn by the user. The contact data is picked up by receiving device 84 located at or near plant 71 when the user is nearby and facing the plant (see dashed arrow 85). The receiving device sends the contact data, together with the URL of the voice service associated with the plant 71, over the home wireless network to the server system 80 and, in particular, to voice browser 3 (see arrow 86). This results in the browser 3 accessing the voice page server 4 to retrieve a first page of the voice service associated with the plant 71. This page (and any subsequent pages) are then interpreted by the voice browser with voice output being passed over the home wireless network to the wireless headset 90 of the user (see arrow 89); voice input from the user 5 is returned over the wireless network to the browser.

15

As with the Figure 7 embodiment, the voice browser could be incorporated in equipment carried by the user.

Variants

Many variants are, of course, possible to the arrangements described above with reference to Figures 7 and 8. For example, rather than using a short-range wireless link to pass the user-related contact data to the receiving device, the latter could be provided with other forms of input means such as a smart card reader, magnetic card reader, keyboard, or even a voice input arrangement (in this case, the captured voice input is supplied to a speech recogniser, generally over the communications infrastructure).

25

In another variant, rather than voice input and output both being effected via the user equipment (mobile entity for the Figure 7 embodiment, wireless headset 90 for the Figure 8 embodiment), voice output or input could be done using local loudspeakers or microphones respectively, connected by the communications infrastructure (for Figure 8, this is the home wireless network though wireline connections are, of course, possible). For

30

example, voice input being done using a microphone carried by the user and voice output done by local loudspeakers.

By having multiple local loudspeakers, and assuming that their locations relative to the plant 71 were known to the voice browser system, the voice browser can control the volume from each speaker to make it appear as if the sound output was coming from the plant. This is particularly useful where there are multiple voice-enabled dumb entities in the same area. A similar effect (making the voice output appear to come from the dumb entity) can also be achieved for users wearing stereo-sound headsets provided the following information is known to the voice browser (or other element responsible for setting output levels between the two stereo channels):

- location of the user relative to the entity (this can be determined in any suitable manner including by using a system such as GPS to accurately position the user, the location of the entity being fixed and known); and
- 15 - the orientation of the user's head (determined, for example, using a magnetic flux compass or solid state gyros incorporated into the headset).

Knowing the user's position or orientation relative to the entity also enables the voice service to be adapted accordingly. For example, a user approaching the back of an entity (typically not a plant) may receive a different voice output from the voice service as compared to a user approaching from the front. Similarly, a user facing away from the entity may be differently spoken to by the entity as compared to a user facing the entity. Also, a user crossing past the entity may be differently spoken to as compared to a user moving directly towards the entity or a user moving directly away from the entity (that is, the voice service is dependent on the user's 'line of approach' –this term here being taken to include line of departure also). The user's position/orientation/line-of-approach relative to the entity can be used to adapt the voice service either on the basis of the user's initial position/orientation/approach to the entity or on an ongoing basis responsive to changes in the user's position/orientation/approach. Information regarding the relative position of the user to the entity does not necessarily require the use of user-location determining technology or magnetic flux compasses or gyroscopes – the simple provision of multiple directional receiving devices can be used to identify the user's position relative to the

entity. Indeed, the beacon devices need not even be directional if they are each located away from the entity along a respective approach route.

Where there are multiple voice-enabled dumb entities in the same area, the equipment
5 carried by the user or the voice browser is preferably arranged to ignore new contact data coming from an entity if the user is still in dialog with another entity (in this respect, end of a dialog can be determined either as a sufficiently long pause by the user, a specific termination command from the user, or a natural end to the voice dialog script). To alleviate any problems with receiving contact data from multiple dumb entities that are
10 close to each other, the short-range transmitter is preferably made highly directional in nature, this being readily achieved where the short-range communication is effected using infrared.

By arranging for the identity of the user to be passed to the voice browser or voice page
15 server, profile data on the user (if available) can be looked up by a database access and used to customise the service to the user.

Other variants are also possible. For example, the user on contacting the voice service can be joined into a session with any other users currently using the voice service in respect of
20 the same entity such that all users at least hear the same voice output of the voice service. This can be achieved by functionality at the voice page server (session management being commonly effected at web page servers) but only to the level of what page is currently served to each user. It is therefore preferred to implement this common session feature at a common voice browser thereby ensuring all users hear the same output at the same time.
25 With respect to voice input by session members, there will generally be a need for the voice service to select one input stream in the case that more than one member speaks at the same time. The selected input voice stream can be relayed to other members by the voice browser to provide an indication as to what input is currently being handled; unselected input is not relayed in this manner.

An extension of this arrangement is to join the user into a session with any other users currently using the voice service in respect of the same local entity and other entities that have been logically associated with that entity, the voice inputs and outputs to and from the voice service being made available to all such users. Thus, if two similar plants that are not
5 located near each other are logically associated, users in dialog with both plants are joined into a common session.

The voice-enabled 'dumb' entity can be provided with associated functionality that is controlled by control data passed from the voice service via the communications
10 infrastructure. This control data is for example, scripted into the voice pages embedded in multimodal tags for extraction by the voice browser and sending to the entity associated functionality (contact data for this functionality having been passed to the voice browser along with the user-related contact data).

15 Where the 'dumb' entity has an associated mouth-like feature movable by associated functionality, the control data from the voice service can be used to cause operation of the mouth-like device in synchronism with voice output from the voice service. Thus a dummy can be made to move its mouth in synchronism with dialog it is uttering via its associated voice service. This feature, which has application in museums and like attractions, is
20 preferably used with the aforementioned arrangement of joining users in dialog with the same entity into a common session – since the dummy can only move its mouth in synchronism with one piece of dialog at a time, having all interested persons in the same session and selecting which user voice input is to be responded to, is clearly advantageous.

25 The mouth-like feature and associated functionality can conveniently be associated with the dumb entity by incorporation into the receiving device and can exist in isolation from any other "living" feature. The mouth-like feature can be either physical in nature with actuators controlling movement of physical parts of the feature, or simply an electronically-displayed mouth (for example displayed on an LCD display). The
30 coordination of the mouth-like feature with the voice service output aids people with hearing difficulties to understand what is being said.

Of course, as well as using multimodal tags for control data to be passed to the entity, more normal multimodal interactions (displays, keyboard, pointing devices etc.) can be scripted in the voice service provided by the voice page server in the embodiments of Figures 7 and

5 8.

CLAIMS

1. A method of voice communication concerning a local entity wherein:
 - 5 (a) - the local entity has an associated voice service hosted on a separate server connected to a communications infrastructure;
 - (b) - upon a user approaching the local entity, contact data relating to the user is transferred to a receiving device that is located at or near the local entity and is connected to the communications infrastructure;
 - 10 (c) - the contact data received by the receiving device is used to establish communication through the communications infrastructure between the voice service and equipment carried by the user that is in wireless connection with the communications infrastructure;
 - (d) - the user interacts with the voice service with the latter acting as voice proxy for the
15 local entity.
2. A method according to claim 1, wherein the contact data is one of:
 - a data connection address for the user's equipment;
 - a telephone number of telephone functionality incorporated into the user's
20 equipment;
 - user-specific data for translation by an element of the communications infrastructure into an access number or address of the user's equipment.
3. A method according to claim 1, wherein step (d) involves voice input by the user and
25 voice output by the service with voice input and voice output being effected by sound input and output devices forming part of the user's equipment.
4. A method according to claim 1, wherein step (c) involves voice input by the user and voice output by the service, voice output being effected using a sound output device
30 forming part of the user's equipment, and voice input being through at least one local sound input device that is associated with the locality of the entity rather than with the user

and is connected with the voice service through the communications infrastructure independently of the user's equipment.

- 5 5. A method according to claim 1, wherein step (c) involves voice input by the user and voice output by the service, voice input being effected using a sound input device forming part of the user's equipment, and voice output being through at least one local sound output device that is associated with the locality of the entity rather than with the user and is connected with the voice service through a communications infrastructure independently of the user's equipment.
- 10 6. A method according to claim 5, wherein sound output is through multiple sound output devices controlled by the voice service so that the sound appears to be originating from said local entity.
- 15 7. A method according to claim 1, wherein the voice service is effected by the serving of voice pages in the form of text with embedded voice markup tags to a voice browser, the voice browser interpreting these pages and carrying out speech recognition of user voice input, text to speech conversion to generate voice output, and dialog management; the voice browser being disposed between a voice page server and the user.
- 20 8. A method according to claim 7, wherein the user-related contact data serves to identify the user and is passed in step (c) directly or indirectly to the voice browser which uses the contact data to look up an access number or address for the user's equipment.
- 25 9. A method according to claim 1, wherein the user equipment includes a mobile phone, step (c) involving placing the voice service and mobile phone in communication.
10. A method according to claim 1, wherein:
- 30 - the voice service is effected by the serving of voice pages in the form of text with embedded voice markup tags to a voice browser, the voice browser interpreting these pages and carrying out speech recognition of user voice input, text to speech

- conversion to generate voice output, and dialog management; the voice browser being disposed between a voice page server and the user; and
- the user equipment includes a mobile phone, step (c) involving placing the voice service and mobile phone in communication.

5

11. A method according to claim 10, wherein the voice browser is not part of the user's equipment and in step (c) the contact data, in the form of information for contacting the user's equipment, is passed directly to the voice browser together with a URL of the voice service, the voice browser contacting the user on the mobile phone using a voice circuit or data connection that is then used in step (d) for voice input and/or output between the user and voice browser.

12. A method according to claim 10, wherein the voice browser is not part of the user's equipment and the contact data comprises user-specific information which the voice browser can use to derive information for contacting the user's equipment, step (c) involving sending the user-specific information to the voice browser together with a URL of the voice service, the voice browser contacting the user on the mobile phone using a voice circuit or data connection that is then used in step (d) for voice input and/or output between the user and voice browser.

20

13. A method according to claim 10, wherein the voice browser is not part of the user's equipment and in step (c) the user-related contact data is passed to the voice page server which is then responsible for passing the contact data to the voice browser, the voice browser using this contact data to contact the user on the mobile phone using a voice circuit or data connection that is then used in step (d) for voice input and/or output between the user and voice browser.

25

14. A method according to claim 10, wherein the voice browser is part of the user's equipment and in step (c) the user-related contact data is passed to the voice page server which then connects with the user equipment via a data-capable bearer service of the communications infrastructure, the data-capable bearer service being subsequently used in

30

step (d) for passing text based input and/or output between the voice browser and voice page server.

15. A method according to any one of claims 1 to 8, wherein the wireless network is a home/office/proprietary-space local network hosting the voice service, the local entity being located in the home/office/proprietary-space concerned.

16. A method according to claim 15, wherein the user equipment includes a wireless headset which in step (d) is used for exchanging voice input and output with the voice service.

17. A method according to claim 1, wherein in step (b) the identity of the user is sent to the voice service and used by the latter to look up user profile data which is then used to customise the voice service to the user.

18. A method according to any one of the preceding claims, wherein the user on being placed in contact with the voice service in step (c) is joined into a session with any other users currently using the voice service in respect of the same local entity such that all users at least hear the voice output of the voice service.

19. A method according to claim 18, wherein voice input from a user is not broadcast to other users joined in the same session unless that input is selected for handling by the voice service.

20. A method according to any one of claims 1 to 17, wherein the user on being placed in contact with the voice service in step (b) is joined into a session with any other users currently using the voice service in respect of the same local entity and other entities that have been logically associated with that entity, the voice inputs and outputs to and from the voice service being made available to all such users.

21. A method according to any one of the preceding claims, wherein the receiving device includes parameter values relating to the state of said local entity in said contact data, these parameter values being passed in step (c) over the communications infrastructure to the voice service where they are used in conditioning the output of the voice service.

5

22. A method according to any one of the preceding claims, wherein the local entity has associated functionality that is controlled by control data passed from the voice service via the communications infrastructure to said functionality.

10 23. A method according to claim 22, wherein the local entity has an associated mouth-like feature movable by said functionality, the control data from the voice service being used to cause operation of the mouth-like feature in synchronism with voice output from the voice service.

15 24. A method according to claim 23, wherein the mouth-like feature is incorporated into the receiving device.

25. A method according to claim 1, wherein the voice service provided to a user is dependent on the user's position, orientation or line of approach relative to the entity.

20

26. A method according to claim 25, wherein multiple receiving devices are associated with the entity, the contact data of the receiving device first or most-recently picking up the user-related contact data determining the voice service being provided to the user in respect of that entity.

25

27. A method according to claim 25, wherein the location of the user is continually monitored and their position relative to the entity is used to determine the voice service provided to the user in respect of that entity.

30

ABSTRACT**Voice Communication Concerning a Local Entity**

5

A local entity (71) without its own means of voice communication is provided with the semblance of having a voice interaction capability. This is done by providing a receiving device (72) at or near the entity, for picking up contact data transmitted by a nearby person wanting to talk to the local entity. This contact data is used by the receiving device (72) to

10 establish communication between a voice service (4) associated with the local entity (71) and equipment carried by the user. The voice service (4) is hosted separately from the local entity, and takes the form, for example, of pages marked up with voice-markup tags for interpretation by a voice browser (3).

15

(Fig. 7)

THIS PAGE BLANK (USPTO)

1 / 4

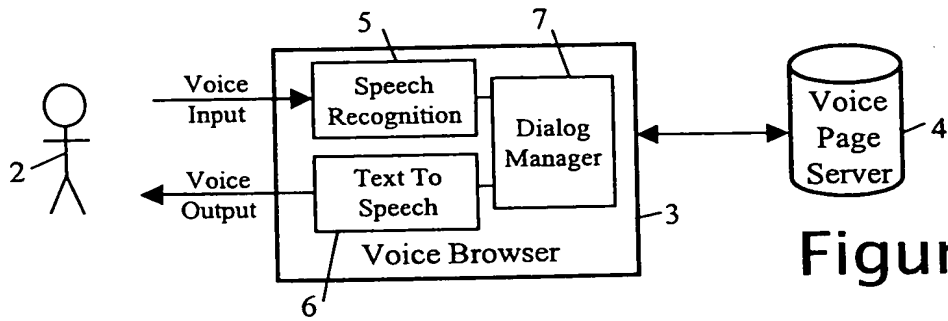


Figure 1

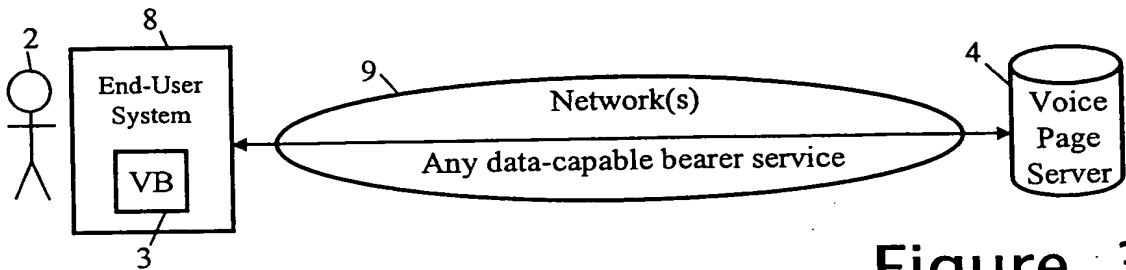


Figure 3

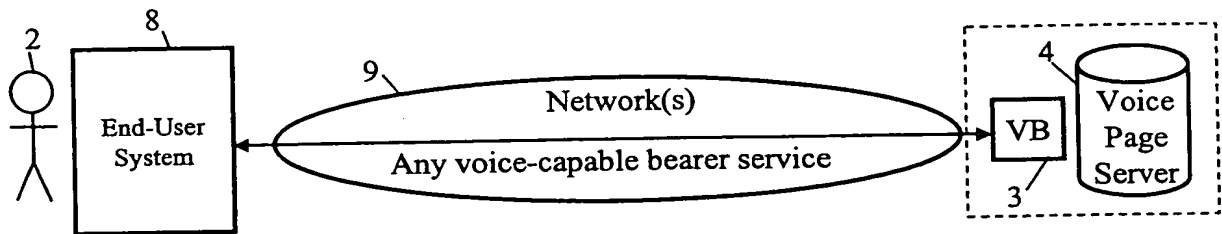


Figure 4

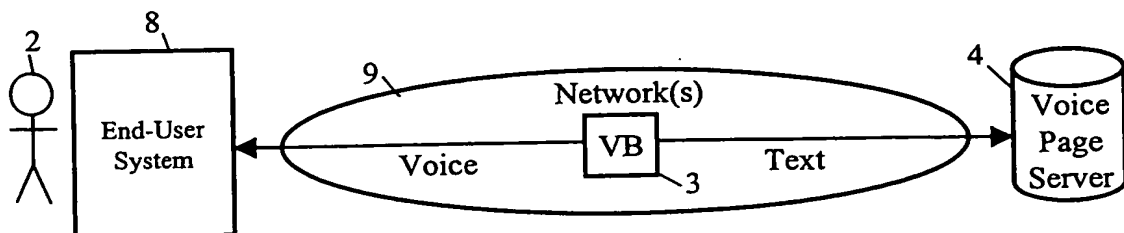


Figure 5

THIS PAGE BLANK (USPTO)

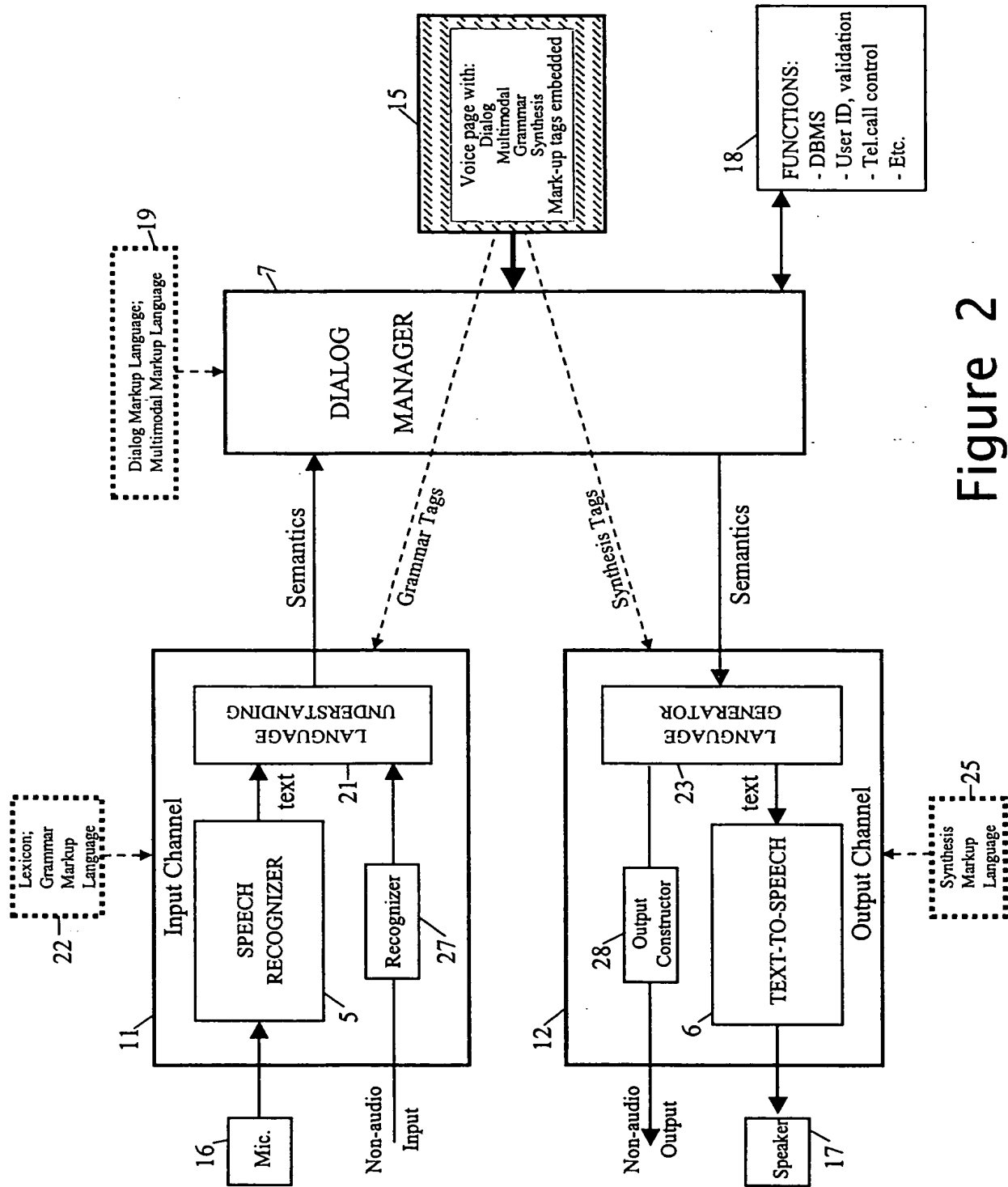


Figure 2

THIS PAGE BLANK (USPTO)

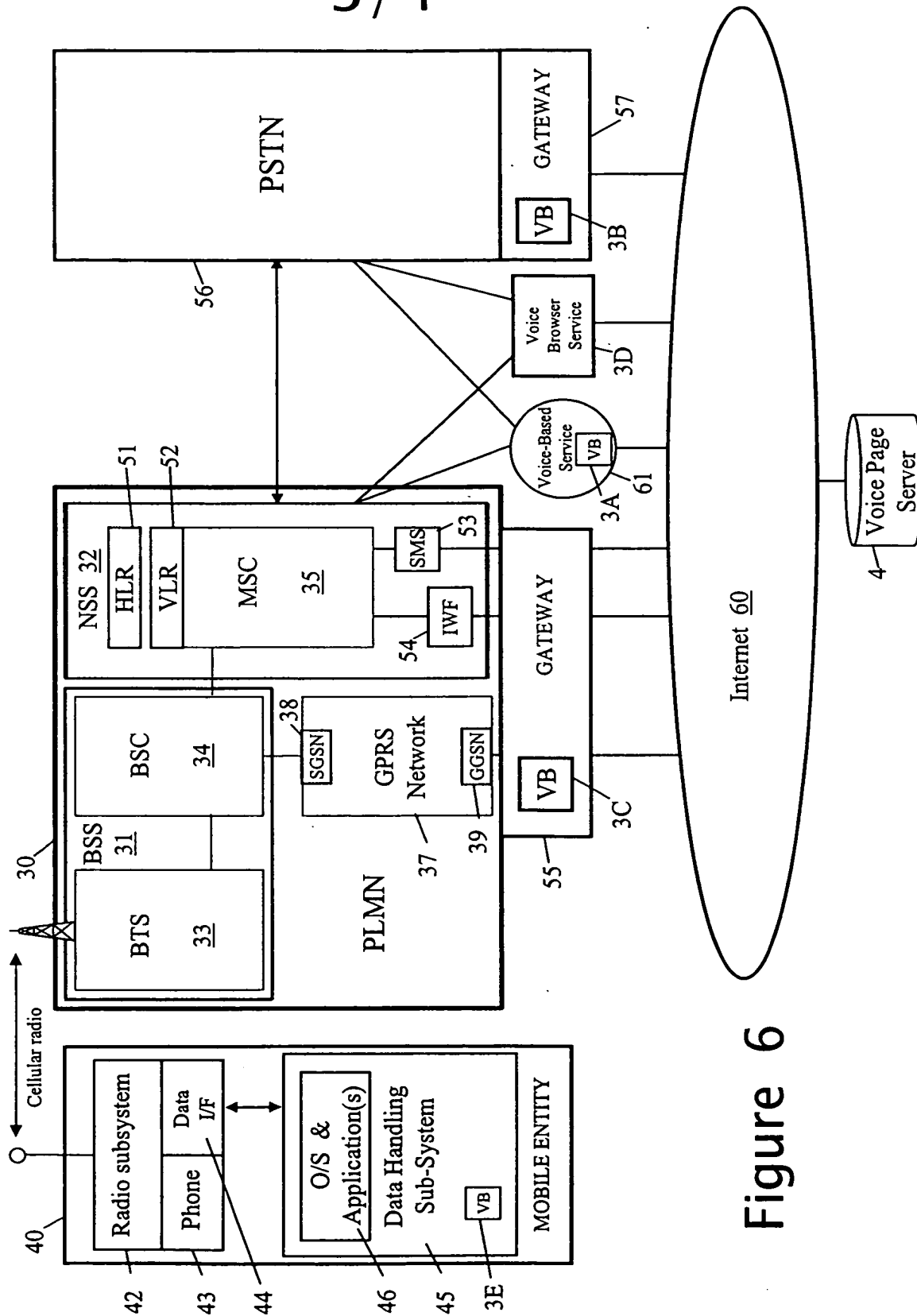


Figure 6

THIS PAGE BLANK (USPTO)

4/4

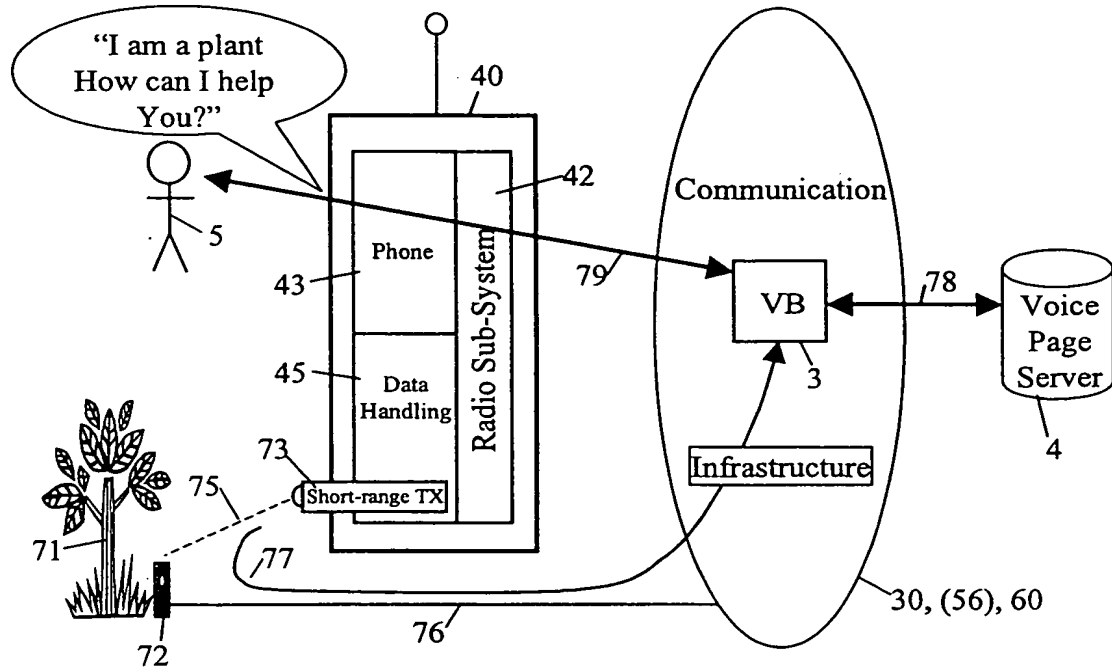


Figure 7

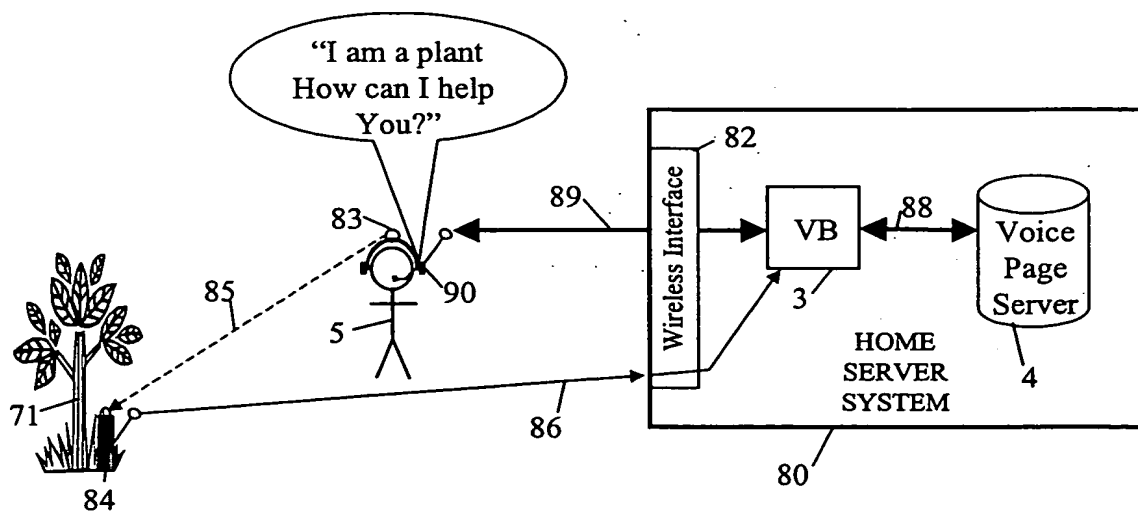


Figure 8

THIS PAGE BLANK (USPTO)